CRYSTALLIZATION AND STRUCTURE DETERMINATION OF *STAPHYLOCOCCUS AUREUS* UDP-N-ACETYLENOLPYRUVYLGLUCOSAMINE REDUCTASE (*S. aureus* MurB)

This application claims the benefit of U.S. Provisional Application Serial No. 60/147,164 filed 4 August 1999, which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

This invention relates to the crystallization and structure determination of *Staphylococcus aureus* UDP-N-acetylenolpyruvylglucosamine reductase (*S. aureus* MurB).

BACKGROUND OF THE INVENTION

Reports of an increase in antibiotic resistant bacteria have stimulated efforts to find new classes of therapeutic agents that will prevent society from entering a "post-antibiotic age." Historically, three important cellular functions have been the major targets of antibiotics - cell wall biosynthesis, DNA replication, and protein translation. The biosynthesis of the bacterial cell wall, in particular the peptidoglycan polymer, is a particularly attractive target since this flexible structure provides protection for the cell against osmotic lysis. To date, most of the therapeutic agents discovered that target cell wall biosynthesis inhibit the later stages of peptidoglycan biosynthesis at the point where interstrand cross linking occurs between the peptide chains. Recent efforts have been directed toward purifying and characterizing all the enzymes in the peptidoglycan biosynthetic pathway with an eye toward designing novel enzyme inhibitors of these essential targets.

Bacterial peptidoglycan is a polymer which includes a repeating disaccharide subunit of *N*-acetylglucosamine and *N*-acetylmuramic acid and an extended four to five residue amino acid chain. The first step toward creating this peptidoglycan polymer involves the formation of UDP-*N*-acetylmuramic acid from UDP-*N*-acetylglucosamine by the enzymes MurA and MurB. MurA

catalyzes the first stage of this transformation by transferring the enolpyruvate moiety of phosphoenolpyruvate to the 3' hydroxyl of UDP-*N*-acetylglucosamine with the release of inorganic phosphate. The resulting product, enolpyruvyl-UDP-*N*-acetylglucosamine (EP-UDPGlcNAc), undergoes a reduction catalyzed by the MurB enzyme by utilizing one equivalent of NADPH and a solvent derived proton. This two electron reduction creates the lactyl ether of UDP- *N*-acetylmuramic acid upon which a five residue peptide chain is built. Construction of this pentapeptide is catalyzed in a nonribosomal fashion by the enzymes MurC, MurD, MurE, and MurF (Figure 1) in both Gram negative bacteria such as *Escherichia coli* and Gram positive bacteria such as *Staphylococcus aureus*. The resulting UDP- *N*-acetylmuramyl pentapeptide is subsequently attached to an undecaprenyl lipid moiety by MraY and joined to another sugar, UDP-*N*-acetylglucosamine by MurG. In *Staphylococci* the next steps of peptidoglycan biosynthesis involve another family of enzymes, FemX, FemA, and FemB which create a pentaglycine strand in a stepwise fashion on the amino terminus of the lysine side chain. This extended Lys-Gly_5 chain serves as the interstrand bridge between nearby peptide strands. Crosslinking between strands can then occur between the lysine-pentapeptide bridge and the carbonyl of the fourth residue (D-Ala) with release of the terminal D-Ala in a transpeptidation step catalyzed by penicillin binding proteins.

While several laboratories have characterized some of the peptidoglycan biosynthetic enzymes for *E. coli,* little biochemistry or structural biology has been carried out on these enzymes in a clinically relevant Gram positive organism. Interest in the molecular mechanisms of peptidoglycan biosynthesis in Gram positive organisms has increased in recent years as methicillin resistant *S. aureus* strains have surfaced that have acquired resistance to the antibiotic vancomycin.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides a method for crystallizing an *S. aureus* MurB molecule or molecular complex that includes preparing purified *S. aureus* MurB at a concentration of about 1 mg/ml to about 50 mg/ml and crystallizing *S. aureus* MurB from a solution comprising about 1

wt. % to about 50 wt. % PEG, 0 wt. % to about 40 wt. % DMSO, about 100 mM to about 1 M ammonium or lithium sulfate, about 0 mM to about 20 mM 2-mercaptoethanol, about 0.005 mM to about 40 mM EP-UDPGlcNAc substrate, and buffered to a pH of about 5 to about 8.

In another aspect, the present invention provides crystalline forms of an *S. aureus* MurB molecule. In one embodiment, a crystal of an *S. aureus* MurB is provided having the trigonal space group symmetry $I2_13$.

In another aspect, the present invention provides a scalable three dimensional configuration of points derived from structure coordinates of at least a portion of an *S. aureus* MurB molecule or molecular complex. In one embodiment, the scalable three dimensional set of points is derived from structure coordinates of at least the backbone atoms of the amino acids representing a FAD and/or substrate binding pocket of an *S. aureus* MurB molecule or molecular complex. In another embodiment, the scalable three dimensional set of points is derived from structure coordinates of at least a portion of a molecule or a molecular complex that is structurally homologous to an *S. aureus* MurB molecule or molecular complex. On a molecular scale, the configuration of points derived from a homologous molecule or molecular complex have a root mean square deviation of less than about 1.0 Å from the structure coordinates of the molecule or complex

In another aspect, the present invention provides a molecule or molecular complex that includes at least a portion of an *S. aureus* MurB FAD and/or substrate binding pocket. In one embodiment, the *S. aureus* MurB FAD binding pocket includes the amino acids listed in Table 1, preferably the amino acids listed in Table 2, and more preferably the amino acids listed in Table 3, the FAD binding pocket being defined by a set of points having a root mean square deviation of less than about 1.7 Å, preferably less than about 1.0 Å, from points representing the backbone atoms of the amino acids. In another embodiment, the *S. aureus* MurB substrate binding pocket includes the amino acids listed in Table 4, preferably the amino acids listed in Table 5, and more preferably the amino acids listed in Table 6, the substrate binding pocket being defined by a set of points having a root mean square deviation of less than about 1.0 Å from points representing the backbone atoms of the amino acids.

TABLE 1

Residues near the FAD binding site in *S. aureus* MurB

Identified residues 4Å away from the FAD

TYR	42	LEU	98	TYR	149	VAL	199
TYR	77	SER	115	MET	150	ARG	225
LEU	78	ILE	140	ALA	152	GLN	229
GLY	79	PRO	141	GLY	153	LEU	231
ASN	80	GLY	142	ALA	154	SER	235
GLY	81	SER	143	ARG	188	GLY	237
SER	82	GLY	145	ILE	192	PHE	274
ASN	83	GLY	146	LEU	197	ARG	310
ILE	84	ALA	147	VAL	198		

TABLE 2

Residues near the FAD binding site in *S. aureus* MurB

Identified residues 7Å away from FAD

THR	41	LEU	99	VAL	148	LEU	200
TYR	42	SER	115	TYR	149	GLU	201
THR	43	GLY	116	MET	150	ARG	225
THR	76	ALA	117	ASN	151	GLN	229
TYR	77	ILE	119	ALA	152	PRO	230
LEU	78	PHE	136	GLY	153	LEU	231
GLY	79	GLY	139	ALA	154	TYR	233
ASN	80	ILE	140	TYR	155	PRO	234
GLY	81	PRO	141	ARG	188	SER	235
SER	82	GLY	142	ILE	192	CYS	236
ASN	83	SER	143	GLN	193	GLY	237
ILE	84	ILE	144	HIS	196	SER	238
ILE	85	GLY	145	LEU	197	PHE	274
ILE	96	GLY	146	VAL	198	ARG	310
LEU	98	ALA	147	VAL	199	ILE	312

TABLE 3

Residues near the FAD binding site in *S. aureus* MurB

Identified residues 10Å away

LEU	37	SER	100	TYR	155	ARG	225
TYR	40	LEU	101	GLY	156	GLU	226
THR	41	ALA	113	GLY	157	LYS	228
TYR	42	GLY	114	GLU	158	GLN	229
THR	43	SER	115	VAL	159	PRO	230
LYS	44	GLY	116	LYS	160	LEU	231
THR	45	ALA	117	ALA	166	GLU	232
TYR	52	ALA	118	LEU	167	TYR	233
PRO	55	ILE	119	CYS	168	PRO	234
VAL	61	ILE	120	VAL	169	SER	235
VAL	65	GLU	135	ASN	170	CYS	236
VAL	75	PHE	136	LEU	183	GLY	237
THR	76	ALA	137	ASP	186	SER	238
TYR	77	CYS	138	TYR	187	VAL	239
LEU	78	GLY	139	ARG	188	SER	268
GLY	79	ILE	140	ASN	189	LYS	270
ASN	80	PRO	141	SER	190	HIS	271
GLY	81	GLY	142	ILE	191	GLY	273
SER	82	SER	143	ILE	192	PHE	274
ASN	83	ILE	144	GLN	193	MET	275
ILE	84	GLY	145	LYS	194	VAL	276
ILE	85	GLY	146	GLU	195	ASN	277
ILE	86	ALA	147	HIS	196	TYR	286
ILE	91	VAL	148	LEU	197	GLU	308
ILE	94	TYR	149	VAL	198	VAL	309
VAL	95	MET	150	VAL	199	ARG	310
ILE	96	ASN	151	LEU	200	ILE	311
SER	97	ALA	152	GLU	201	ILE	312
LEU	98	GLY	153	ALA	202		
LEU	99	ALA	154	LEU	221		

TABLE 4

Residues near the EP-UDPGlcNAc binding site in *S. aureus* MurB

Identified residues 4Å away from EP-UDPGlcNAc

TYR	155	GLN	229	GLN	253	PHE	274
TYR	187	GLY	237	GLN	258	FAD	401
ARG	188	SER	238	HIS	271		
ARG	225	LYS	250	ALA	272		

TABLE 5

Residues near the EP-UDPGlcNAc binding site in *S. aureus* MurB

Identified residues 7Å away from EP-UDPGlcNAc

SER	82	ARG	188	ARG	242	SER	268
ASN	83	ARG	224	PHE	247	THR	269
GLY	139	ARG	225	ALA	248	LYS	270
ILE	140	LYS	228	GLY	249	HIS	271
PRO	141	GLN	229	LYS	250	ALA	272
MET	150	CYS	236	LEU	251	GLY	273
GLY	153	GLY	237	ILE	252	PHE	274
ALA	154	SER	238	GLN	253	GLU	308
TYR	155	VAL	239	ASP	254	FAD	401
GLY	156	PHE	240	GLN	258		
TYR	187	GLN	241	VAL	267		

TABLE 6

Residues near the EP-UDPGlcNAc binding site in *S. aureus* MurB

Identified residues 10Å away from EP-UDPGlcNAc

TYR	42	TYR	155	SER	238	GLN	258
THR	43	GLY	156	VAL	239	GLY	259
GLY	81	GLY	157	PHE	240	VAL	267
SER	82	GLU	158	GLN	241	SER	268
ASN	83	TYR	187	ARG	242	TMR	269
ILE	84	ARG	188	PRO	243	LYS	270
PHE	136	LEU	221	HIS	246	HIS	271
ALA	137	ARG	224	PHE	247	ALA	272
CYS	138	ARG	225	ALA	248	GLY	273
GLY	139	GLU	226	GLY	249	PHE	274
ILE	140	SER	227	LYS	250	MET	275
PRO	141	LYS	228	LEU	251	ASN	306
GLY	142	GLN	229	ILE	252	ARG	307
MET	150	PRO	230	GLN	253	GLU	308
ASN	151	LEU	231	ASP	254	VAL	309
ALA	152	SER	235	SER	255	FAD	401
GLY	153	CYS	236	ASN	256		
ALA	154	GLY	237	LEU	257		

In another aspect, the present invention provides molecules or molecular complexes that are structurally homologous to an *S. aureus* MurB molecule or molecular complex.

In another aspect, the present invention provides a machine readable storage medium including the structure coordinates of all or a portion of an *S. aureus* MurB molecule, molecular complex, a structurally homologous

molecule or complex, including structurally equivalent structures, as defined herein, particularly an FAD or substrate binding pocket thereof, or a similarly shaped homologous binding pocket. A storage medium encoded with these data is capable of displaying on a computer screen, or similar viewing device, a three-dimensional graphical representation of a molecule or molecular complex which comprises a binding pocket or a similarly shaped homologous binding pocket.

In another aspect, the present invention provides a method for identifying inhibitors, ligands, and the like of an *S. aureus* MurB molecule by providing the coordinates of a molecule of *S. aureus* MurB to a computerized modeling system; identifying chemical entities that are likely to bind to or interfere with the molecule (e.g., screening a small molecule library); and, optionally, procuring or synthesizing and assaying the compounds or analogues derived therefrom for bioactivity. In another aspect, the present invention provides methods for designing inhibitors, ligands, and the like by providing the coordinates of a molecule of *S. aureus* MurB to a computerized modeling system; designing a chemical entity that is likely to bind to or interfere with the molecule; and, optionally, synthesizing the chemical entity and assaying the chemical entity for bioactivity. In another aspect, the present invention provides inhibitors and ligands designed by the above method. In one embodiment, a composition is provided that includes an inhibitor or ligand designed or identified by the above method. In another embodiment, the composition is a pharmaceutical composition.

In another aspect, the present invention provides a method involving molecular replacement to obtain structural information about a molecule or molecular complex of unknown structure. The method includes crystallizing the molecule or molecular complex, generating an x-ray diffraction pattern from the crystallized molecule or molecular complex, and applying at least a portion of the structure coordinates set forth in Figure 4 to the x-ray diffraction pattern to generate a three-dimensional electron density map of at least a portion of the molecule or molecular complex.

In another aspect, the present invention provides a method for homology modeling an *S. aureus* MurB homolog.

DEFINITIONS

Two crystallographic data sets (with structure factors F) are considered isomorphous if, after scaling,

$$\frac{\Delta F}{F} = \frac{\Sigma |F_1 - F_2|}{\Sigma F_1}$$

is less than about 35% for the reflections between 8 Å and 4 Å.

ABBREVIATIONS

The following abbreviations are used throughout this disclosure:

UDP-*N*-acetylenolpyruvylglucosamine reductase (MurB).

Uridine diphospho-*N*-acetylglucosamine (UDPGlcNAc).

Uridine diphospho-*N*-acetylglucosamine enolpyruvate (EP-UDPGlcNAc).

Uridine diphospho-*N*-acetylmuramic acid (UDPMurNAc).

Reduced β-nicotinamide adenine dinucleotide phosphate (NADPH).

Isopropylthio-β-D-galactoside (IPTG).

Dithiothreitol (DTT).

Flavin adenine dinucleotide (FAD).

Dimethyl sulfoxide (DMSO).

Multiple anomalous dispersion (MAD).

The following amino acid abbreviations are used throughout this disclosure:

A = Ala = Alanine	T = Thr = Threonine
V = Val = Valine	C = Cys = Cysteine
L = Leu = Leucine	Y = Tyr = Tyrosine
I = Ile = Isoleucine	N = Asn = Asparagine
P = Pro = Proline	Q = Gln = Glutamine
F = Phe = Phenylalanine	D = Asp = Aspartic Acid
W = Trp = Tryptophan	E = Glu = Glutamic Acid
M = Met = Methionine	K = Lys = Lysine
G = Gly = Glycine	R = Arg = Arginine
S = Ser = Serine	H = His = Histidine

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the pathway for the biosynthesis of the UDP-*N*-acetylmuramyl pentapeptide, detailing the first two steps catalyzed by the enzymes MurA and MurB.

Figure 2 shows an amino acid sequence alignment for recombinant *S. aureus* (SEQ ID NO:1, which includes the His_6 region) and *E. coli* MurB. Dots in the sequences indicate gaps inserted in order to optimize the alignment. Identical residues are indicate by | and similar residues are indicated by . and : symbols. Sequence alignment was performed using the program GAP (GCG Version 9, Genetics Computational Group, Madison, WI). Residues corresponding to the FAD binding region (domains 1 and 2) are overlined and those corresponding to the substrate binding region (domain 3) are dash underlined. Structural features that are present in the *E. coli* but not the *S. aureus* protein are boxed and were omitted from superpositions. In regions where there are significant deletions in the *S. aureus* protein compared to the *E. coli* protein, no significance should be given to the placement of the connecting

residues in the *S. aureus* sequence. Protein sequences not observed due to disordered electron density for the N and C termini of *S. aureus* MurB are underlined.

Figure 3 shows a) solvent flattened MAD electron density map at 2.3 Å resolution for residues 110-115 with the final model and b) final $2F_o$-F_c electron density map at 2.3 Å for residues 110-115 with the final model.

Figure 4 lists the atomic structure coordinates for molecule *S. aureus* MurB as derived by x-ray diffraction from a crystal of that complex. The following abbreviations are used in Figure 4:

"Atom" refers to the element whose coordinates are measured. The second column defines the number of the atom in the structure. The letters in the third column define the element. The fourth and fifth columns define the amino acid and the number of the amino acid in the structure, respectively.

"X, Y, Z" crystallographically define the atomic position of the element measured.

"Occ" is an occupancy factor that refers to the fraction of the molecules in which each atom occupies the position specified by the coordinates. A value of "1" indicates that each atom has the same conformation, i.e., the same position, in all molecules of the crystal.

"B" is a thermal factor that measures movement of the atom around its atomic center.

Figure 5 shows ribbon diagrams of a) *S. aureus* MurB structure with bound FAD cofactor and b) *E. coli* MurB structure with bound FAD cofactor and EP-UDPGlcNAc substrate.

Figure 6 shows secondary structure diagram for a) *S. aureus* MurB and b) *E. coli* MurB. The domain assignments are indicated with a dotted line. Naming of the secondary structure was made to correspond to the previously published *E. coli* MurB structure. Where new elements of secondary structure are present in the *S. aureus* structure, naming includes an "A."

Figure 7 shows the superposition of *S. aureus* MurB (thick lines) and *E. coli* MurB (thin lines). Residues used for superpositions of the various domains are: a) All C_α atoms in common between the two structures. b) C_α atoms in domain 1 (lower right of molecule as shown in part a). The arrow

points to the additional N-terminal α helix and β strand present in the *S. aureus* MurB structure. c) C_α atoms in domain 2 (upper center of molecule as shown in part a). The arrow points to the Tyr 190 loop in the *E. coli* MurB structure which is absent in the *S. aureus* MurB structure. d) C_α atoms in domain 3 (lower left of molecule as shown in part a). The arrow points to the single split βαββ fold in the *E. coli* MurB structure that is absent in the *S. aureus* MurB structure.

Figure 8 shows a schematic view of side chain and main chain interactions between *S. aureus* MurB and the FAD cofactor. Residues for the *S. aureus* enzyme ("SA" prefix) are indicated adjacent to the amino acid along with the corresponding residues for the *E. coli* enzyme ("EC" prefix). Distances which would allow hydrogen bonds with the FAD are shown as dotted lines.

Figure 9 shows proposed binding interactions of *S. aureus* MurB with the EP-UDPGlcNAc substrate ("SA prefix and parentheses) based on the *E. coli* MurB structure. Residues involved in EP-UDPGlcNAc binding in *E.* coli MurB are also shown ("EC" prefix).

Figure 10 shows the sequence alignment of a representative sample of MurB sequences from Genbank: MURB_HELPY (*Helicobacter pylori*), MURB_AQUAE (*Aquifex aeolicus*), MURB_BACSU (*Bacillus subtilis*), MURB_BORBU (*Borrelia burgdorferi*), MURB_CHLPN (*Chlamydia pneumoniae*), MURB_RICPR (*Rickettsia prowazekii*), MURB_SAURE (*Staphylococcus aureus*), MURB_ECOLI (*Escherichia coli*), MURB_HAEIN (*Haemophilus influenzae*), MURB_SALTY (*Salmonella typhimurium*), and MURB_BORDE (*Bordetella pertussis*). Asterisks indicate the proposed active site residues involved in catalysis. Several other MurB sequences were not included in this alignment including *Treponema pallidum* MurB (class II MurB) and *Mycobacterium tuberculosis* MurB (class I MurB) because of additional insertions or deletions in these protein sequences which complicated the multiple sequence alignment.

Figure 11 shows C_α backbone traces from three MurB structures (superposition based on the flavin binding domains). The two *E. coli* MurB structures (substrate free *E. coli* MurB shown as dotted lines, EP-UDPGlcNAc bound MurB shown as thin lines) show that the substrate binding domain is actually closer to the flavin binding domain in the absence of substrate than

when the substrate is bound. The *S. aureus* MurB structure reveals the greatest displacement of the substrate binding domain from the flavin domain of the three structures.

Figure 12 lists the structure factors and multiple anomalous dispersion phases for the crystal structure of *S. aureus* MurB. "INDE" refers to the indices h, k, and l (columns 2, 3, and 4 respectively) of the lattice planes. "FOBS" refers to the structure factor of the observed reflections. "SIGMA" is the standard deviation for the observations. "PHAS" refers to the phase used for the observations. "FOM" refers to the figure of merit.

DETAILED DESCRIPTION OF THE INVENTION

Crystalline Form(s) and Method of Making

Applicants have produced crystals comprising *S. aureus* MurB that are suitable for x-ray crystallographic analysis. The three-dimensional structure of *S. aureus* MurB was solved using high resolution x-ray crystallography. Preferably, the crystal has the cubic space group $I2_13$. More preferably, the crystal comprises cubic shaped unit cells, each unit cell having the dimensions $a = b = c = 178.9 \pm 20$ Å with $a = \beta = \gamma = 90°$. The crystallized enzyme has one molecule in the asymmetric unit and includes a bound FAD cofactor.

Purified *S. aureus* MurB, preferably at a concentration of about 1 mg/ml to about 50 mg/ml, may be crystallized, for example, using the sitting or hanging drop procedure from a solution including about 1 wt. % to about 50 wt. % polyethylene glycol (PEG, preferably having a number average molecular weight between about 200 and about 20,000), 0 to about 40 wt. % DMSO, about 100 mM to about 1 M ammonium or lithium sulfate, about 0 mM to about 20 mM 2-mercaptoethanol, about 0.005 mM to about 40 mM EP-UDPGlcNAc substrate, and buffered to a pH of about 5 to about 8. Use of a buffer having a pK_a of between about 4 and 9 is preferred. Variation in buffer and buffer pH as well as other additives such as PEG is apparent to those skilled in the art and may result in similar crystals.

Accordingly, one embodiment of the invention provides an *S. aureus* MurB or *S. aureus* MurB/ligand crystal.

The invention further includes an *S. aureus* MurB crystal or *S. aureus* MurB/ligand crystal that is isomorphous with an *S. aureus* MurB crystal characterized by a unit cell having the dimensions a = b = c = 178.9 ± 20 Å with $a = \beta = \gamma = 90°$.

X-ray Crystallographic Analysis

Crystals of *S. aureus* MurB in the cubic space group $I2_13$ with cell constants a=b=c=178.9 Å, $\alpha=\beta=\gamma=90°$ diffracted to 2.3 Å resolution. Initial attempts with molecular replacement using the *E. coli* MurB coordinates (T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)) were unsuccessful despite the near 50% similarity with the *S. aureus* sequence (Figure 2). Therefore, an independent set of phases was derived using multiple anomalous dispersion (MAD) with selenomethionine incorporated protein. *S. aureus* selenomethionine MurB was prepared by inhibiting endogenous methionine biosynthesis while supplementing the expressing cells with selenomethionine (G.D. Van Duyne et al., J. Mol. Biol. 229, 105-24 (1993); T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)). Methionine biosynthesis down regulation eliminates the need for transferring the protein expression vector into a *met*⁻ strain. This technique reduces the time and effort required for producing selenomethionine incorporated protein and results in near quantitative incorporation of selenomethionine into the overexpressed protein. Anomalous and dispersive difference Pattersons revealed the presence of four selenium sites. Solvent-flattened multiple anomalous dispersion phases to 2.3 Å revealed an exceptionally clear electron density map with no significant breaks in the main chain. A portion of the electron density map is shown in Figure 3. The structure was refined to 2.3 Å resolution with an R-factor of 20.3% and a Free R-factor of 22.3% as described in Tables 7and 8.

Table 7. Data collection and phasing statistics

	λ 1.0332 Å (12,000 eV)	λ 0.97939 Å (12,659.4 eV)	λ 0.97928 Å (12,660.8 eV)
Resolution	2.3 Å	2.3 Å	2.3 Å
No. observations	252,156	267,578	268,391
No. unique refl.	39,984	40,336	40,394
% completeness	94.4%	95.2%	95.3%
R_{sym}	7.5%	9.5%	9.4%
R_{cullis} acentrics	--	0.77	0.83
R_{cullis} anomalous	0.99	0.84	0.84
Phasing power			
Centrics	--	0.87	0.69
acentrics	--	0.77	0.83
Mean figure of merit (to 2.3 Å resolution)			
before solvent flattening		0.464	
after solvent flattening		0.605	

Table 8. Refinement Statistics

	R-factor	Free R-factor	No. of reflections
10-2.3 Å F≥ 2σ	20.3%	22.3%	33,156
		Bonds (Å)	Angles(°)
r.m.s deviation from ideal geometry		0.008	1.37

	Number of atoms	Average B-factor
Protein	2345	28.4
Waters	213	36.6
FAD	53	23.6
Total	2611	29.0

Each of the constituent amino acids of *S. aureus* MurB is defined by a set of structure coordinates as set forth in Figure 4. The term "structure coordinates" refers to Cartesian coordinates derived from mathematical equations related to the patterns obtained on diffraction of a monochromatic beam of x-rays by the atoms (scattering centers) of an *S. aureus* MurB complex in crystal form. The diffraction data are used to calculate an electron density map of the repeating unit of the crystal. The electron density maps are then used to establish the positions of the individual atoms of the *S. aureus* MurB protein or protein/ligand complex.

Slight variations in structure coordinates can be generated by mathematically manipulating the *S. aureus* MurB or *S. aureus* MurB/ligand structure coordinates. For example, the structure coordinates set forth in Figure 4 could be manipulated by crystallographic permutations of the structure coordinates, fractionalization of the structure coordinates, integer additions or subtractions to sets of the structure coordinates, inversion of the structure coordinates or any combination of the above. Alternatively, modifications in the crystal structure due to mutations, additions, substitutions, and/or deletions of amino acids, or other changes in any of the components that make up the crystal, could also yield variations in structure coordinates. Such slight variations in the

individual coordinates will have little effect on overall shape. If such variations are within an acceptable standard error as compared to the original coordinates, the resulting three-dimensional shape is considered to be structurally equivalent. Structural equivalence is described in more detail below.

It should be noted that slight variations in individual structure coordinates of the *S. aureus* MurB or *S. aureus* MurB/ligand complex, as defined above, would not be expected to significantly alter the nature of chemical entities such as ligands that could associate with the binding pockets. In this context, the phrase "associating with" refers to a condition of proximity between a chemical entity, or portions thereof, and an *S. aureus* thymidylate kinase molecule or portions thereof. The association may be non-covalent, wherein the juxtaposition is energetically favored by hydrogen bonding, van der Waals forces, or electrostatic interactions, or it may be covalent. Thus, for example, a ligand that bound to or interfered with the active site binding pocket of *S. aureus* MurB would also be expected to bind to or interfere with another binding pocket whose structure coordinates define a shape that falls within the acceptable error.

It will be readily apparent to those of skill in the art that the numbering of amino acids in other isoforms of *S. aureus* thymidylate kinase may be different than that of *S. aureus* thymidylate kinase expressed in *E. coli*.

Overview of the Structure

S. aureus MurB is composed of three domains (Figures 5 and 6). Domains 1 and 2 are responsible for binding of the flavin adenine dinucleotide (FAD) cofactor while domain 3 is responsible for substrate binding. The r.m.s. deviation for all C_α atoms in common between the *E. coli* and *S. aureus* structures (236 residues out of the 326 *S. aureus* residues) is 2.20 Å (Figure 7a). Superpositions for each of the three domains in *S. aureus* compared to their respective domains in the *E. coli* enzyme resulted in slightly better superpositions for domains 2 and 3 (Figure 7b-d). Domain 1 (residues 14-98) of *S. aureus* MurB has a r.m.s. deviation of 2.20 Å for the C_α atoms compared to *E. coli* MurB. The second domain of *S. aureus* MurB (residues 101-229) has a r.m.s. deviation of 1.80 Å for the C_α atoms of the corresponding residues in *E. coli* MurB. The r.m.s. deviation for domain 3 of the *S. aureus* enzyme (residues

230-316) is 1.05 Å for the C_α atoms corresponding to the portion of this domain present in the *E. coli* structure.

While the overall fold of the *S. aureus* MurB enzyme is similar to that of *E. coli* MurB, several exceptions indicate that the *S. aureus* MurB structure represents a significant structural variation for the UDP-*N*-acetylenolpyruvylglucosamine reductases. The first major difference is the additional 32 amino acids that are present at the N-terminus of *S. aureus* MurB which are not present in the *E. coli* enzyme. While only 18 of these amino acids are observed in the electron density map (the remaining 14 amino acids are disordered), these residues form an additional α helix (αA) and β strand (βA) at the beginning of the peptide chain. Similar secondary structure elements in *S. aureus* MurB have been given the corresponding names from *E. coli* MurB (T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)) in order to facilitate the discussion and new elements of secondary structure have been assigned lettered names. The βA strand adds an antiparallel strand to the central parallel β barrel which forms the base of domain 1. This β barrel has an extremely hydrophobic core burying the side chains of residues Leu 37, Leu 78, Ile 84, Ile 86, Ile 91, Val 95, and Ile 312.

The second major structural difference in *S. aureus* MurB is the absence of the loop between β13 and α3 (residues 183 to 203 of *E. coli* MurB). To compensate for the loss of this loop, an additional turn of helix is added to α3 in order to make the connection between β13 and α3 in *S. aureus* MurB. This loop serves an important role in *E. coli* MurB by positioning Tyr 190 to interact directly with the α phosphate of the EP-UDPGlcNAc ligand and to close the active site upon substrate binding (T.E. Benson et al., Structure 4, 47-54 (1996)). Observations of the *S. aureus* MurB structure do not reveal any direct substitutes for Tyr 190 suggesting that this specific mechanism for substrate binding observed in *E. coli* MurB is not utilized in the *S. aureus* enzyme. In the *E. coli* MurB structure, the α phosphate of the EP-UDPGlcNAc also interacts with the side chain of Lys 217. This residue is conserved in the *S. aureus* MurB structure as residue Lys 228. Therefore, one would expect that the absence of Tyr 190 would lead to an increased importance for Lys 228 in the formation of a productive enzyme-substrate complex in *S. aureus* MurB. The third major

structural difference is the deletion of a portion of the substrate binding domain in the *S. aureus* MurB protein structure. One of the α+β motifs present in the *E. coli* structure, the single split βαββ fold (β14, α4, β15, β16) is absent in the *S. aureus* MurB structure. This portion of the *E. coli* structure provides several van der Waals contacts with the EP-UDPGlcNAc ligand. In the absence of this portion of the substrate binding domain, the ligand binding surface on *S. aureus* MurB is notably more narrow.

Several minor differences in the main chain between the two structures are also observed. First, in the *E. coli* MurB, two residues in β5 and β6 are not found in the *S. aureus* MurB structure resulting in two shorter β strands for the *S. aureus* enzyme. Second, the hydrogen bonding distances and geometries for residues 143-151 in *S. aureus* MurB are consistent with a secondary structure assignment of an α helix (αB) for these residues in the core of the protein. This same region in the *E. coli* structure (residues 113-121) shows secondary structure similar to an α helix with allowed phi and psi angles, but with poor hydrogen bonding distances and geometries; therefore, this region was not assigned as an α helix in the original *E. coli* MurB structure (T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)). Third, a single residue deletion in the *S. aureus* structure occurs in the loop between β9 and β10 leading to a shorter connection between these two strands. Finally, the last beta strand (β21) and alpha helix (α7) in the *E. coli* MurB structure are not observed in the *S. aureus* MurB structure. The exact secondary structure of these residues is unknown since the last nine residues of the C terminus of *S. aureus* MurB are disordered in the electron density map.

Flavin Cofactor Binding

The conserved protein fold of MurB in the *S. aureus* enzyme is particularly pronounced for the flavin binding portion of the molecule (domains 1 and 2 - Figures 7 b,c). Not only is the flavin binding fold conserved, but also the conformation of the flavin adenine dinucleotide ligand is nearly identical between the two structures (r.m.s. deviation for all of the cofactor atoms comparing the *S. aureus* FAD and the *E. coli* FAD is 0.30 Å). This FAD binding fold in both MurB structures is a member of a new superfamily of flavin

adenine dinucleotide binding proteins (A.G. Murzin, Cur. Op. Struct. Biol. 6, 386-94 (1996)). The other members of this FAD binding protein family for which protein structures have been solved include *p*-cresol methylhydroxylase from *Pseudomonas putida* (F.S. Matthews et al., Biochemistry 30, 238-47 (1991)), vanillyl-alcohol oxidase from *Penicllium simplicissimum* (A. Mattevi et al., Structure 5, 907-20 (1997)), and CO dehydrogenase from *Oligotropha carboxidovorans* (H. Dobbek et al., Proc. Natl. Acad. Sci USA 96, 8884-89 (1999)).

The protein side chain and main chain interactions with the FAD cofactor found in the *S. aureus* enzyme are mostly similar to those interactions observed in *E. coli* MurB. The N5 and O4 of the isoalloxazine ring of the flavin adenine dinucleotide interact with the guanidinium moiety of Arg 225 in a manner similar to that observed for Arg 214 in the *E. coli* MurB structure (Fig. 8). The sequence and positional conservation of this arginine in the structure suggests that it plays a role in the binding of the flavin and stabilization of the reduced cofactor during catalysis. Two other interactions to the isoalloxazine ring (N3 and O2) are maintained by the main chain nitrogen and carbonyl oxygen of Gly 153. Interactions with the ribityl sugar moiety are also similar to those observed in the *E. coli* MurB. The carbonyl oxygen of Pro 141 and the hydroxyl group of Ser 82 make hydrogen bonds to the first hydroxyl group and the carbonyl of Gly 146 interacts with the third hydroxyl moiety. The extensive interactions between the protein and the diphosphoadenine portion of the molecule are also well conserved with the main chain atoms of residues 79-83 which include the Gly-X-Gly motif found in the Rossman fold and provide important stabilizing interactions with the two phosphates. Two serines (Ser 82 and Ser 143) again serve to make specific contacts with the β phosphate moiety, but using a geometry different from *E. coli* MurB. The placement of Ser 82 is conserved with respect to its counterpart in *E. coli* MurB (Ser 50), but Ser 143 is three residues away (one turn of α helix B) from the position corresponding to the *E. coli* MurB Ser 116. While this places the serine hydroxyl on the opposite site of the phosphate (when compared to the *E. coli* MurB), the hydrogen bonding interaction is maintained. Finally, two residues make contacts to the 3' hydroxyl of the ribityl sugar, Tyr 149 and Arg 310. The corresponding residues

in *E. coli* MurB for Tyr 149 is Ile 149 which does not make any hydrogen bonds to the ribityl sugar, but does make hydrophobic interactions with this part of the flavin. The substitution of tyrosine in *S. aureus* MurB at this position allows for both hydrogen bonding and hydrophobic contacts.

Active Site Arrangement and Implications for Substrate Binding

The electron density map in the active site of *S. aureus* MurB reveals regions of disconnected electron density that did not clearly resemble the EP-UDPGlcNAc substrate despite the presence of the substrate in the crystallization conditions. Therefore, the exact location and interactions between *S. aureus* MurB and the EP-UDPGlcNAc and/or NADPH substrate in the active site cannot be determined at this time. Analysis of the packing in the crystal lattice reveals that a symmetry related molecule protrudes into the active site of *S. aureus* MurB in this crystal. Specifically, two strands of the major beta sheet (β5 and β6) are situated at the active site opening. Superimposing the *E. coli*-EP-UDPGlcNAc bound structure on the *S. aureus* MurB structure indicates that these strands from the symmetry related molecule would interfere with the binding of the uridine portion of the substrate. Many attempts were made to obtain a ligand bound crystal form of *S. aureus* MurB, but no crystallization conditions were identified which would support both crystallization of the protein and binding of EP-UDPGlcNAc.

Although this crystal form of *S. aureus* MurB does not contain either of the MurB substrates EP-UDPGlcNAc or NADPH, comparison of the *S. aureus* and *E. coli* structures reveals strict conservation of the active site residues. The three catalytic active site residues in *E. coli*, Arg 159 and Glu 325 that would stabilize the C2 acicarbanionic species and Ser 229 that would provide a proton to quench the intermediate, are strictly conserved in the *S. aureus* active site - Ser 238, Arg 188, and Glu 308 (Figure 9) (T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)). In addition, two residues that were shown to play a role in substrate binding in the *E. coli* enzyme are also strictly conserved in the active site of *S. aureus* MurB (Tyr 187 and Lys 288). Two other residues shown to be important for substrate binding - Asn 233 and Gln 288 - are replaced in the *S. aureus* structure by Arg 242 and His 271

respectively. The only critical substrate binding residue that is not found in the *S. aureus* structure is the *E. coli* Tyr 190 which is absent due to the deletion of the loop between β13 and α3 in *S. aureus* protein as previously noted. The absence of this single residue does not appear to compromise the ability of the enzyme to bind substrate since all of the other hydrogen bond interactions observed in the *E. coli* MurB structure would be maintained.

A comparison of the two MurB structures reveals that the *S. aureus* MurB possesses the same general fold that was observed in the *E. coli* MurB structure – an α + β protein with three domains. Two of these domains create a binding site for the flavin adenine dinucleotide cofactor and the third domain participates in substrate binding. Based on the high similarity between the two protein sequences, a related fold was clearly expected. Yet it is not surprising that differences between the two structures are evident in the substrate binding regions of the enzymes, since the sequence alignment reveals regions of significant deletions. The most relevant deletions that occur in the *S. aureus* MurB structure involve portions of the enzyme that play an important part in the binding of EP-UDPGlcNAc in *E. coli* MurB. The loop between β13 and α3 in *E. coli* MurB contains Tyr 190 which undergoes a dramatic motion upon substrate binding. The absence of this loop and its key residue in the *S. aureus* structure suggests that ligand binding in this bacterial species has adapted to compensate for the loss of Tyr 190. The K_m for EP-UDPGlcNAc with *S. aureus* MurB is 15μM (S. Swaney, personal communication) which is of the same magnitude as the K_m for EP-UDPGlcNAc with *E. coli* MurB. This kinetic parameter suggests that the loss of Tyr 190 has little impact on the enzyme's ability to bind substrate. It is also possible that charged residues from other parts of the molecule could play a role in ligand binding. Interestingly, the other significant deletion also involves a part of the enzyme involved in substrate binding – the single split βαββ fold (β14, α4, β15 and β16 in *E. coli* MurB). This portion of protein structure has a less specific purpose in the mechanism of action of MurB, but does serve to provide a surface on which the uridine portion of the EP-UDPGlcNAc substrate rests.

Although the MurB crystals were grown in the presence of the substrate EP-UDPGlcNAc, no interpretable electron density for the substrate was

observed. The failure to obtain a substrate complex could be either the result of weak substrate binding to the enzyme under the conditions necessary for crystallization or the result of crystallization conditions which select for a crystal form that excludes substrate. Monitoring ligand binding by UV-visible spectroscopy shows the expected red shift of the flavin absorption spectrum associated with EP-UDPGlcNAc as observed with the *E. coli* MurB enzyme (T.E. Benson et al., Biochemistry 36, 796-805 (1997)) suggesting that the substrate should be bound to the oxidized form of the *S. aureus* MurB. When the initial crystals did not show the presence of substrate in the active site, higher concentrations of substrate were employed in an attempt to produce a substrate complex. Since the original crystallization buffer was at pH 6.5, crystals were also grown at pH 8.0 within the optimal pH range for the *S. aureus* MurB enzymatic activity. Unfortunately, neither of these changes resulted in formation of a substrate complex in the crystals that were obtained. These results strongly suggest that the crystal form which is favored during crystallization selects against preserving the *S. aureus* MurB-EP-UDPGlcNAc complex.

Placement of the EP-UDPGlcNAc substrate can be inferred from the *E. coli* MurB-EP-UDPGlcNAc structure. The active site for MurB is defined by residues which stabilize the intermediate produced when a hydride is transferred from N1 to C3 of the enolpyruvyl group and facilitate the quenching of this intermediate. Based on *S. aureus* MurB structure, a mechanism was proposed for stabilization of this acicarbanionic species by charge neutralization and/or by protonation. As in the *E. coli* model, two residues, Arg 188 and Glu 308, exist which could serve this mechanistic purpose. Similarly, Ser 238 corresponds to *E. coli* MurB Ser 229 which is proposed to serve as a general acid catalyst in conjunction with an active site water (T.E. Benson et al., Biochemistry 36, 806-11 (1997)). The Ser 238 hydroxyl is 6.3Å away from N1 of the FAD cofactor in the *S. aureus* MurB structure. Since the corresponding serine hydroxyl is similarly positioned (6.1 Å from N1) in the *E. coli* MurB structure, it appears that *S. aureus* MurB is well-positioned for reduction of the enolpyruvyl group.

An analysis of available MurB sequences in Genbank provides increasing evidence that these three active site residues play critical roles in

stabilization and reduction of the C2 acicarbanion intermediate during catalysis. In the sequence alignment shown in Figure 10, the active site glutamate (Glu 308 in *S. aureus* MurB and Glu 325 in *E. coli* MurB) is strictly conserved. The active site arginine (Arg 188 in *S. aureus* MurB and Arg 159 in *E. coli* MurB) is conserved in all species with the exception of *Borrelia burgdorfei* MurB in which a conservative substitution to a lysine is made. The active site serine is equally well conserved across species with the only exception being the MurB sequence from *Chlamydia pneumoniae* where a conservative Cys substitution is present. This conservation of active site residues suggests that mechanistically the UDP-*N*-acetylenolpyruvylglucosamine reductases are equivalent with respect to the reduction of the enolpyruvyl group.

In contrast, these sequence alignments also suggest that there are at least two structural scaffolds for the MurB family. Comparison of the sequences from Genbank reveals that the *S. aureus* MurB is not the only MurB sequence in which the *E. coli* Tyr 190 loop and the single split βαββ fold from the substrate binding domain (β14, α4, β15 and β16 in *E. coli* MurB) are absent (Figure 10). This striking similarity among sequences indicates that many of these MurBs from other species would adopt the *S. aureus* MurB type three-dimensional structure as opposed to the *E. coli* MurB type three-dimensional structure. Thus, the trend suggests that at least two distinct classes of MurB protein structures exist – that of the *E. coli* type (type I MurB) and that of the *S. aureus* type (type II MurB) – which distinguishes the construction of the substrate binding domain and the manner in which these enzymes bind their substrates.

This classification of two types of MurB consists not only of secondary structural elements that are present or absent, but also of how these structural elements define a mode of substrate binding. A comparison of the substrate domain positions between *S. aureus* MurB and the two forms of *E. coli* MurB (substrate free and EP-UDPGlcNAc bound forms) is shown in Figure 11. This superposition reveals that *S. aureus* MurB's substrate binding domain is notably more displaced from domains 1 and 2 in either of the substrate free or the EP-UDPGlcNAc bound forms of *E. coli* MurB. While interactions between crystallographically related molecules hold domain 3 in this open conformation,

this displacement reveals the flexibility of the enzyme to achieve an open conformation of the *S. aureus* enzyme with ready access to the ligand binding site. The process of binding ligand in the *S. aureus* MurB most likely involves closing of the enzyme by direct movement of domain 3 towards the flavin binding domains. In contrast, *E. coli* MurB facilitates ligand binding by an indirect mechanism involving the movement of domain 3 away from the flavin binding domain leading to the disruption of the stacking interaction between Tyr 190 and Tyr 254. This motion allows Tyr 190 to adopt a new rotameric configuration which provides a hydrogen bond to the α phosphate and closes off the active site from the solvent.

Binding pockets/active sites/other structural features

Applicants' invention has provided, for the first time, information about the shape and structure of the cofactor and substrate binding pockets of *S. aureus* MurB.

Binding pockets are of significant utility in fields such as drug discovery. The association of natural ligands or substrates with the binding pockets of their corresponding receptors or enzymes is the basis of many biological mechanisms of action. Similarly, many drugs exert their biological effects through association with the binding pockets of receptors and enzymes. Such associations may occur with all or any parts of the binding pocket. An understanding of such associations helps lead to the design of drugs having more favorable associations with their target, and thus improved biological effects. Therefore, this information is valuable in designing potential inhibitors of *S. aureus* MurB-like binding pockets, as discussed in more detail below.

A "molecular complex" means a protein in covalent or non-covalent association with a chemical entity or compound. The term "binding pocket" or "active site" as used herein, refers to a region of a molecule or molecular complex, that, as a result of its shape, favorably associates with another chemical entity. Thus, a binding pocket may include or consist of features such as cavities, surfaces, or interfaces between domains. Chemical entities that may associate with a binding pocket include, but are not limited to, cofactors, substrates, inhibitors, agonists, antagonists, etc.

The FAD binding pocket of *S. aureus* MurB is located on the interface between domains 1 and 2, and preferably includes the amino acids listed in Table 1, more preferably the amino acids listed in Table 2, and most preferably the amino acids listed in Table 3, as represented by the structure coordinates listed in Figure 4. It will be readily apparent to those of skill in the art that the numbering of amino acids in other isoforms of *S. aureus* MurB may be different than that of recombinant *S. aureus* MurB expressed in *E. coli*. Alternatively, the FAD binding pocket of *S. aureus* MurB includes those amino acids whose backbone atoms are situated within about 4 Å, more preferably within about 7 Å, most preferably within about 10 Å, of one or more constituent atoms of a bound FAD cofactor or analog, as determined from the structure coordinates listed in Figure 4. Alternatively, the FAD binding pocket comprises those amino acids whose backbone atoms are situated within a sphere centered on the coordinates representing the alpha carbon atom of residue Asn 80, the sphere having a radius of about 16 Å, preferably about 20 Å, and more preferably about 25 Å.

The substrate binding pocket of *S. aureus* MurB is located in domain 3, and preferably includes the amino acids listed in Table 4, more preferably the amino acids listed in Table 5, and most preferably the amino acids listed in Table 6, as represented by the structure coordinates listed in Figure 4. Alternatively, the substrate binding pocket of *S. aureus* MurB includes those amino acids whose backbone atoms are situated within about 4 Å, more preferably within about 7 Å, most preferably within about 10 Å, of one or more constituent atoms of a bound substrate or inhibitor, as determined from the structure coordinates listed in Figure 4. Alternatively, the substrate binding pocket comprises those amino acids whose backbone atoms are situated within a sphere centered on the coordinates representing the alpha carbon atom of residue Ser 238, the sphere having a radius of about 12 Å, preferably about 20 Å, and more preferably about 25 Å.

The amino acid constituents of an *S. aureus* MurB binding pocket as defined herein, as well as selected constituent atoms thereof, are positioned in three dimensions in accordance with the structure coordinates listed in Figure 4. In one aspect, the structure coordinates defining the binding pocket of *S. aureus*

MurB include structure coordinates of all atoms in the constituent amino acids; in another aspect, the structure coordinates of the binding pocket include structure coordinates of just the backbone atoms of the constituent atoms.

The term "*S. aureus* MurB-like binding pocket" refers to a portion of a molecule or molecular complex whose shape is sufficiently similar to at least a portion of a cofactor or substrate binding pocket of *S. aureus* MurB as to be expected to bind a structurally related cofactor such as FAD or structurally related substrates such EP-UDPGlcNAc and/or NADPH. A structurally equivalent binding pocket is defined by a root mean square deviation from the structure coordinates of the backbone atoms of the amino acids that make up the binding pockets in *S. aureus* MurB (as set forth in Figure 4) of at most about 1.5 Å. How this calculation is obtained is described below.

Accordingly, the invention thus provides molecules or molecular complexes comprising an *S. aureus* MurB binding pocket or *S. aureus* MurB-like binding pocket, as defined by the sets of structure coordinates described above.

Three-Dimensional Configurations

X-ray structure coordinates define a unique configuration of points in space. Those of skill in the art understand that a set of structure coordinates for protein or an protein/ligand complex, or a portion thereof, define a relative set of points that, in turn, define a configuration in three dimensions. A similar or identical configuration can be defined by an entirely different set of coordinates, provided the distances and angles between coordinates remain essentially the same. In addition, a scalable configuration of points can be defined by increasing or decreasing the distances between coordinates by a scalar factor while keeping the angles essentially the same.

The present invention thus includes the scalable three-dimensional configuration of points derived from the structure coordinates of at least a portion of an *S. aureus* MurB molecule or molecular complex, as listed in Figure 4, as well as structurally equivalent configurations, as described below. Preferably, the scalable three-dimensional configuration includes points derived

from structure coordinates representing the locations of a plurality of the amino acids defining an *S. aureus* MurB binding pocket.

In one embodiment, the scalable three-dimensional configuration includes points derived from structure coordinates representing the locations of the backbone atoms of a plurality of amino acids defining the *S. aureus* MurB FAD binding pocket, preferably the amino acids listed is Table 1, more preferably the amino acids listed in Table 2, and most preferably the amino acids listed in Table 3; in another embodiment, the three-dimensional configuration includes points derived from structure coordinates representing the locations of the side chain and the backbone atoms (other than hydrogens) of a plurality of the amino acids defining the *S. aureus* MurB FAD binding pocket, preferably the amino acids listed is Table 1, more preferably the amino acids listed in Table 2, and most preferably the amino acids listed in Table 3.

In another embodiment, the scalable three-dimensional configuration includes points derived from structure coordinates representing the locations of the backbone atoms of a plurality of amino acids defining the *S. aureus* MurB substrate binding pocket, preferably the amino acids listed is Table 4, more preferably the amino acids listed in Table 5, and most preferably the amino acids listed in Table 6; in another embodiment, the three-dimensional configuration includes points derived from structure coordinates representing the locations of the side chain and the backbone atoms (other than hydrogens) of a plurality of the amino acids defining the *S. aureus* MurB substrate binding pocket, preferably the amino acids listed is Table 4, more preferably the amino acids listed in Table 5, and most preferably the amino acids listed in Table 6.

Likewise, the invention also includes the scalable three-dimensional configuration of points derived from structure coordinates of molecules or molecular complexes that are structurally homologous to *S. aureus* MurB, as well as structurally equivalent configurations. Structurally homologous molecules or molecular complexes are defined below. Advantageously, structurally homologous molecules can be identified using the structure coordinates of *S. aureus* MurB (Figure 4) according to a method of the invention.

The configurations of points in space derived from structure coordinates according to the invention can be visualized as, for example, a holographic image, a stereodiagram, a model or a computer-displayed image, and the invention thus includes such images, diagrams or models.

Structurally Equivalent Crystal Structures

Various computational analyses can be used to determine whether a molecule or the binding pocket portion thereof is "structurally equivalent," defined in terms of its three-dimensional structure, to all or part of *S. aureus* MurB or its binding pockets. Such analyses may be carried out in current software applications, such as the Molecular Similarity application of QUANTA (Molecular Simulations Inc., San Diego, CA) version 4.1, and as described in the accompanying User's Guide.

The Molecular Similarity application permits comparisons between different structures, different conformations of the same structure, and different parts of the same structure. The procedure used in Molecular Similarity to compare structures is divided into four steps: (1) load the structures to be compared; (2) define the atom equivalences in these structures; (3) perform a fitting operation; and (4) analyze the results.

Each structure is identified by a name. One structure is identified as the target (i.e., the fixed structure); all remaining structures are working structures (i.e., moving structures). Since atom equivalency within QUANTA is defined by user input, for the purpose of this invention equivalent atoms are defined as protein backbone atoms (N, Cα, C, and O) for all conserved residues between the two structures being compared. A conserved residue is defined as a residue that is structurally or functionally equivalent. Only rigid fitting operations are considered.

When a rigid fitting method is used, the working structure is translated and rotated to obtain an optimum fit with the target structure. The fitting operation uses an algorithm that computes the optimum translation and rotation to be applied to the moving structure, such that the root mean square difference of the fit over the specified pairs of equivalent atom is an absolute minimum. This number, given in angstroms, is reported by QUANTA.

For the purpose of this invention, any molecule or molecular complex or binding pocket thereof, or any portion thereof, that has a root mean square deviation of conserved residue backbone atoms (N, Cα, C, O) of less than about 1.5 Å, when superimposed on the relevant backbone atoms described by the reference structure coordinates listed in Figure 4, is considered structurally equivalent" to the reference molecule. That is to say, the crystal structures of those portions of the two molecules are substantially identical, within acceptable error. Particularly preferred structurally equivalent molecules or molecular complexes are those that are defined by the entire set of structure coordinates listed in Figure 4, ± a root mean square deviation from the conserved backbone atoms of those amino acids of not more than 1.5 Å. More preferably, the root mean square deviation is less than about 1.0 Å.

The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations. It is a way to express the deviation or variation from a trend or object. For purposes of this invention, the "root mean square deviation" defines the variation in the backbone of a protein from the backbone of *S. aureus* MurB or a binding pocket portion thereof, as defined by the structure coordinates of *S. aureus* MurB described herein.

Machine Readable Storage Media

Transformation of the structure coordinates for all or a portion of *S. aureus* MurB or the *S. aureus* MurB/ligand complex or one of its binding pockets, for structurally homologous molecules as defined below, or for the structural equivalents of any of these molecules or molecular complexes as defined above, into three-dimensional graphical representations of the molecule or complex can be conveniently achieved through the use of commercially-available software.

The invention thus further provides a machine-readable storage medium comprising a data storage material encoded with machine readable data which, when using a machine programmed with instructions for using said data, is capable of displaying a graphical three-dimensional representation of any of the molecule or molecular complexes of this invention that have been described above. In a preferred embodiment, the machine-readable data storage medium

comprises a data storage material encoded with machine readable data which, when using a machine programmed with instructions for using said data, is capable of displaying a graphical three-dimensional representation of a molecule or molecular complex comprising all or any parts of an *S. aureus* MurB binding pocket or an *S. aureus* MurB-like binding pocket, as defined above. In another preferred embodiment, the machine-readable data storage medium is capable of displaying a graphical three-dimensional representation of a molecule or molecular complex defined by the structure coordinates of all of the amino acids listed in Figure 4, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5 Å.

In an alternative embodiment, the machine-readable data storage medium comprises a data storage material encoded with a first set of machine readable data which comprises the Fourier transform of the structure coordinates set forth in Figure 4, and which, when using a machine programmed with instructions for using said data, can be combined with a second set of machine readable data comprising the x-ray diffraction pattern of a molecule or molecular complex to determine at least a portion of the structure coordinates corresponding to the second set of machine readable data.

For example, a system for reading a data storage medium may include a computer comprising a central processing unit ("CPU"), a working memory which may be, e.g., RAM (random access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more display devices (e.g., cathode-ray tube ("CRT") displays, light emitting diode ("LED") displays, liquid crystal displays ("LCDs"), electroluminescent displays, vacuum fluorescent displays, field emission displays ("FEDs"), plasma displays, projection panels, etc.), one or more user input devices (e.g., keyboards, microphones, mice, touch screens, etc.), one or more input lines, and one or more output lines, all of which are interconnected by a conventional bidirectional system bus. The system may be a stand-alone computer, or may be networked (e.g., through local area networks, wide area networks, intranets, extranets, or the internet) to other systems (e.g., computers, hosts, servers, etc.). The system may also include additional computer controlled devices such as consumer electronics and appliances.

Input hardware may be coupled to the computer by input lines and may be implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may comprise CD-ROM drives or disk drives. In conjunction with a display terminal, a keyboard may also be used as an input device.

Output hardware may be coupled to the computer by output lines and may similarly be implemented by conventional devices. By way of example, the output hardware may include a display device for displaying a graphical representation of a binding pocket of this invention using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

In operation, a CPU coordinates the use of the various input and output devices, coordinates data accesses from mass storage devices, accesses to and from working memory, and determines the sequence of data processing steps. A number of programs may be used to process the machine-readable data of this invention. Such programs are discussed in reference to the computational methods of drug discovery as described herein. References to components of the hardware system are included as appropriate throughout the following description of the data storage medium.

Machine-readable storage devices useful in the present invention include, but are not limited to, magnetic devices, electrical devices, optical devices, and combinations thereof. Examples of such data storage devices include, but are not limited to, hard disk devices, CD devices, digital video disk devices, floppy disk devices, removable hard disk devices, magneto-optic disk devices, magnetic tape devices, flash memory devices, bubble memory devices, holographic storage devices, and any other mass storage peripheral device. It should be understood that these storage devices include necessary hardware (e.g., drives, controllers, power supplies, etc.) as well as any necessary media (e.g., disks, flash cards, etc.) to enable the storage of data.

Structurally Homologous Molecules, Molecular Complexes, And Crystal Structures

The structure coordinates set forth in Figure 4 can be used to aid in obtaining structural information about another crystallized molecule or molecular complex. The method of the invention allows determination of at least a portion of the three-dimensional structure of molecules or molecular complexes which contain one or more structural features that are similar to structural features of *S. aureus* MurB. These molecules are referred to herein as "structurally homologous" to *S. aureus* MurB. Similar structural features can include, for example, regions of amino acid identity, conserved active site or binding site motifs, and similarly arranged secondary structural elements (e.g., α helices and β sheets). Optionally, structural homology is determined by aligning the residues of the two amino acid sequences to optimize the number of identical amino acids along the lengths of their sequences; gaps in either or both sequences are permitted in making the alignment in order to optimize the number of identical amino acids, although the amino acids in each sequence must nonetheless remain in their proper order. Preferably, two amino acid sequences are compared using the Blastp program, version 2.0.9, of the BLAST 2 search algorithm, as described by Tatiana et al., FEMS Microbiol Lett 174, 247-50 (1999), and available at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. Preferably, the default values for all BLAST 2 search parameters are used, including matrix = BLOSUM62; open gap penalty = 11, extension gap penalty = 1, gap x_dropoff = 50, expect = 10, wordsize = 3, and filter on. In the comparison of two amino acid sequences using the BLAST search algorithm, structural similarity is referred to as identity." Preferably, a structurally homologous molecule is a protein that has an amino acid sequence sharing at least 65% identity with a native or recombinant amino acid sequence of *S. aureus* MurB (for example, SEQ ID NO:1). More preferably, a protein that is structurally homologous to *S. aureus* MurB includes at least one contiguous stretch of at least 50 amino acids that shares at least 80% amino acid sequence identity with the analogous portion of the native or recombinant *S. aureus* MurB (for example, SEQ ID NO:1). Methods for generating structural information about the structurally homologous

molecule or molecular complex are well-known and include, for example, molecular replacement techniques.

Therefore, in another embodiment this invention provides a method of utilizing molecular replacement to obtain structural information about a molecule or molecular complex whose structure is unknown comprising the steps of:

(a) crystallizing the molecule or molecular complex of unknown structure;

(b) generating an x-ray diffraction pattern from said crystallized molecule or molecular complex; and

(c) applying at least a portion of the structure coordinates set forth in Figure 4 to the x-ray diffraction pattern to generate a three-dimensional electron density map of the molecule or molecular complex whose structure is unknown.

By using molecular replacement, all or part of the structure coordinates of *S. aureus* MurB or the *S. aureus* MurB/ligand complex as provided by this invention (and set forth in Figure 4) can be used to determine the structure of a crystallized molecule or molecular complex whose structure is unknown more quickly and efficiently than attempting to determine such information *ab initio*.

Molecular replacement provides an accurate estimation of the phases for an unknown structure. Phases are a factor in equations used to solve crystal structures that cannot be determined directly. Obtaining accurate values for the phases, by methods other than molecular replacement, is a time-consuming process that involves iterative cycles of approximations and refinements and greatly hinders the solution of crystal structures. However, when the crystal structure of a protein containing at least a structurally homologous portion has been solved, the phases from the known structure provide a satisfactory estimate of the phases for the unknown structure.

Thus, this method involves generating a preliminary model of a molecule or molecular complex whose structure coordinates are unknown, by orienting and positioning the relevant portion of *S. aureus* MurB or the *S. aureus* MurB/ligand complex according to the structure coordinates listed in Figure 4 within the unit cell of the crystal of the unknown molecule or molecular complex

so as best to account for the observed x-ray diffraction pattern of the crystal of the molecule or molecular complex whose structure is unknown. Phases can then be calculated from this model and combined with the observed x-ray diffraction pattern amplitudes to generate an electron density map of the structure whose coordinates are unknown. This, in turn, can be subjected to any well-known model building and structure refinement techniques to provide a final, accurate structure of the unknown crystallized molecule or molecular complex (E. Lattman, "Use of the Rotation and Translation Functions," in Meth. Enzymol., 115, pp. 55-77 (1985); M.G. Rossman, ed., "The Molecular Replacement Method," Int. Sci. Rev. Ser., No. 13, Gordon & Breach, New York (1972)).

Structural information about a portion of any crystallized molecule or molecular complex that is sufficiently structurally homologous to a portion of *S. aureus* MurB can be resolved by this method. In addition to a molecule that shares one or more structural features with *S. aureus* MurB as described above, a molecule that has similar bioactivity, such as the same catalytic activity, substrate specificity or ligand binding activity as *S. aureus* MurB, may also be sufficiently structurally homologous to *S. aureus* MurB to permit use of the structure coordinates of *S. aureus* MurB to solve its crystal structure.

In a preferred embodiment, the method of molecular replacement is utilized to obtain structural information about a molecule or molecular complex, wherein the molecule or molecular complex comprises at least one *S. aureus* MurB subunit or homolog. A "subunit" of *S. aureus* MurB is an *S. aureus* MurB molecule that has been truncated at the N-terminus or the C-terminus, or both. In the context of the present invention, a "homolog" of *S. aureus* MurB is a protein that contains one or more amino acid substitutions, deletions, additions, or rearrangements with respect to the amino acid sequence of *S. aureus* MurB, but that, when folded into its native conformation, exhibits or is reasonably expected to exhibit at least a portion of the tertiary (three-dimensional) structure of *S. aureus* MurB. For example, structurally homologous molecules can contain deletions or additions of one or more contiguous or noncontiguous amino acids, such as a loop or a domain.

Structurally homologous molecules also include "modified" *S. aureus* MurB molecules that have been chemically or enzymatically derivatized at one or more constituent amino acid, including side chain modifications, backbone modifications, and N- and C- terminal modifications including acetylation, hydroxylation, methylation, amidation, and the attachment of carbohydrate or lipid moieties, cofactors, and the like.

A heavy atom derivative of *S. aureus* MurB is also included as an *S. aureus* MurB homolog. The term "heavy atom derivative" refers to derivatives of *S. aureus* MurB produced by chemically modifying a crystal of *S. aureus* MurB. In practice, a crystal is soaked in a solution containing heavy metal atom salts, or organometallic compounds, e.g., lead chloride, gold thiomalate, thiomersal or uranyl acetate, which can diffuse through the crystal and bind to the surface of the protein. The location(s) of the bound heavy metal atom(s) can be determined by x-ray diffraction analysis of the soaked crystal. This information, in turn, is used to generate the phase information used to construct three-dimensional structure of the protein (T.L. Blundell and N.L. Johnson, Protein Crystallography, Academic Press (1976)).

Because it is expected that *S. aureus* MurB can crystallize in more than one crystal form, the structure coordinates of *S. aureus* MurB as provided by this invention are particularly useful in solving the structure of other crystal forms of *S. aureus* MurB or *S. aureus* MurB complexes.

The structure coordinates of *S. aureus* MurB as provided by this invention are particularly useful in solving the structure of *S. aureus* MurB mutants. Mutants may be prepared, for example, by expression of *S. aureus* MurB cDNA previously altered in its coding sequence by oligonucleotide-directed mutagenesis. Mutants may also be generated by site-specific incorporation of unnatural amino acids into MurB proteins using the general biosynthetic method of C.J. Noren et al., Science, 244:182-188 (1989). In this method, the codon encoding the amino acid of interest in wild-type *S. aureus* MurB is replaced by a "blank" nonsense codon, TAG, using oligonucleotide-directed mutagenesis. A suppressor tRNA directed against this codon is then chemically aminoacylated *in vitro* with the desired unnatural amino acid. The aminoacylated tRNA is then added to an *in vitro* translation system to yield a

mutant *S. aureus* thymidylate kinase with the site-specific incorporated unnatural amino acid.

Selenocysteine or selenomethionine may be incorporated into wild-type or mutant *S. aureus* MurB by expression of *S. aureus* MurB-encoding cDNAs in auxotrophic *E. coli* strains (W.A. Hendrickson et al., EMBO J., 9(5):1665-1672 (1990)). In this method, the wild-type or mutagenized *S. aureus* MurB cDNA may be expressed in a host organism on a growth medium depleted of either natural cysteine or methionine (or both) but enriched in selenocysteine or selenomethionine (or both). Alternatively, selenomethionine analogues may be prepared by down regulation methionine biosynthesis. (T.E. Benson et al., Nat. Struct. Biol., 2:644-53 (1995); G.D. Van Duyne et al., J. Mol. Biol. 229:105-24 (1993)).

The structure coordinates of *S. aureus* MurB listed in Figure 4 are also particularly useful to solve the structure of crystals of *S. aureus* MurB, *S. aureus* MurB mutants or *S. aureus* MurB homologs co-complexed with a variety of chemical entities. This approach enables the determination of the optimal sites for interaction between chemical entities, including candidate *S. aureus* MurB inhibitors and *S. aureus* MurB. Potential sites for modification within the various binding site of the molecule can also be identified. This information provides an additional tool for determining the most efficient binding interactions, for example, increased hydrophobic interactions, between *S. aureus* MurB and a chemical entity. For example, high resolution x-ray diffraction data collected from crystals exposed to different types of solvent allows the determination of where each type of solvent molecule resides. Small molecules that bind tightly to those sites can then be designed and synthesized and tested for their *S. aureus* MurB inhibition activity.

All of the complexes referred to above may be studied using well-known x-ray diffraction techniques and may be refined versus 1.5-3 Å resolution x-ray data to an R value of about 0.20 or less using computer software, such as X-PLOR (Yale University, 81992, distributed by Molecular Simulations, Inc.; see, e.g., Blundell & Johnson, supra; Meth. Enzymol., Vol. 114 & 115, H.W. Wyckoff et al., eds., Academic Press (1985)). This information may thus be

used to optimize known *S. aureus* MurB inhibitors, and more importantly, to design new *S. aureus* MurB inhibitors.

The invention also includes the unique three-dimensional configuration defined by a set of points defined by the structure coordinates for a molecule or molecular complex structurally homologous to *S. aureus* MurB as determined using the method of the present invention, structurally equivalent configurations, and magnetic storage media comprising such set of structure coordinates.

Further, the invention includes structurally homologous molecules as identified using the method of the invention.

Homology Modeling

Using homology modeling, a computer model of an *S. aureus* MurB homolog can be built or refined without crystallizing the homolog. First, a preliminary model of the *S. aureus* MurB homolog is created by sequence alignment with *S. aureus* MurB, secondary structure prediction, the screening of structural libraries, or any combination of those techniques. Computational software may be used to carry out the sequence alignments and the secondary structure predictions. Structural incoherences, e.g., structural fragments around insertions and deletions, can be modeled by screening a structural library for peptides of the desired length and with a suitable conformation. For prediction of the side chain conformation, a side chain rotamer library may be employed. Where the *S. aureus* MurB homolog has been crystallized, the final homology model can be used to solve the crystal structure of the homolog by molecular replacement, as described above. Next, the preliminary model is subjected to energy minimization to yield an energy minimized model. The energy minimized model may contain regions where stereochemistry restraints are violated, in which case such regions are remodeled to obtain a final homology model. The homology model is positioned according to the results of molecular replacement, and subjected to further refinement comprising molecular dynamics calculations.

Rational Drug Design

Computational techniques can be used to screen, identify, select and design chemical entities capable of associating with *S. aureus* MurB or structurally homologous molecules. Knowledge of the structure coordinates for *S. aureus* MurB permits the design and/or identification of synthetic compounds and/or other molecules which have a shape complementary to the conformation of an *S. aureus* MurB binding site. In particular, computational techniques can be used to identify or design chemical entities, such as inhibitors, agonists and antagonists, that associate with an *S. aureus* MurB binding pocket or an *S. aureus* MurB-like binding pocket. Inhibitors may bind to or interfere with all or a portion of the active site of *S. aureus* MurB, and can be competitive, non-competitive, or uncompetitive inhibitors. Once identified and screened for biological activity, these inhibitors/agonists/antagonists may be used therapeutically or prophylactically to block *S. aureus* MurB activity and, thus, inhibit the growth of the bacteria or cause its death. Structure-activity data for analogs of ligands that bind to or interfere with *S. aureus* MurB or *S. aureus* MurB-like binding pockets can also be obtained computationally.

The term "chemical entity," as used herein, refers to chemical compounds, complexes of two or more chemical compounds, and fragments of such compounds or complexes. Chemical entities that are determined to associate with *S. aureus* MurB are potential drug candidates.

Data stored in a machine-readable storage medium that is capable of displaying a graphical three-dimensional representation of the structure of *S. aureus* MurB or a structurally homologous molecule, as identified herein, or portions thereof may thus be advantageously used for drug discovery. The structure coordinates of the chemical entity are used to generate a three-dimensional image that can be computationally fit to the three-dimensional image of *S. aureus* MurB or a structurally homologous molecule. The three-dimensional molecular structure encoded by the data in the data storage medium can then be computationally evaluated for its ability to associate with chemical entities. When the molecular structures encoded by the data is displayed in a graphical three-dimensional representation on a computer screen, the protein

structure can also be visually inspected for potential association with chemical entities.

One embodiment of the method of drug design involves evaluating the potential association of a known chemical entity with *S. aureus* MurB or a structurally homologous molecule, particularly with an *S. aureus* MurB binding pocket (e.g., an FAD binding pocket, a substrate binding pocket, etc.) or *S. aureus* MurB-like binding pocket. The method of drug design thus includes computationally evaluating the potential of a selected chemical entity to associate with any of the molecules or molecular complexes set forth above. This method comprises the steps of: (a) employing computational means to perform a fitting operation between the selected chemical entity and a binding pocket, or a pocket nearby the substrate binding pocket, of the molecule or molecular complex; and (b) analyzing the results of said fitting operation to quantify the association between the chemical entity and the binding pocket.

In another embodiment, the method of drug design involves computer-assisted design of chemical entities that associate with *S. aureus* MurB, its homologs, or portions thereof. Chemical entities can be designed in a step-wise fashion, one fragment at a time, or may be designed as a whole or "*de novo*."

To be a viable drug candidate, the chemical entity identified or designed according to the method must be capable of structurally associating with at least part of an *S. aureus* MurB or *S. aureus* MurB-like binding pockets, and must be able, sterically and energetically, to assume a conformation that allows it to associate with the *S. aureus* MurB or *S. aureus* MurB-like binding pocket. Non-covalent molecular interactions important in this association include hydrogen bonding, van der Waals interactions, hydrophobic interactions, and electrostatic interactions. Conformational considerations include the overall three-dimensional structure and orientation of the chemical entity in relation to the binding pocket, and the spacing between various functional groups of an entity that directly interact with the *S. aureus* MurB-like binding pocket or homologs thereof.

Optionally, the potential binding of a chemical entity to an *S. aureus* MurB or *S. aureus* MurB-like binding pocket is analyzed using computer

modeling techniques prior to the actual synthesis and testing of the chemical entity. If these computational experiments suggest insufficient interaction and association between it and the *S. aureus* MurB or *S. aureus* MurB-like binding pocket, testing of the entity is obviated. However, if computer modeling indicates a strong interaction, the molecule may then be synthesized and tested for its ability to bind to or interfere with an *S. aureus* MurB or *S. aureus* MurB-like binding pocket. Binding assays to determine if a compound actually binds to *S. aureus* MurB can also be performed and are well known in the art. Binding assays may employ kinetic or thermodynamic methodology using a wide variety of techniques including, but not limited to, microcalorimetry, circular dichroism, capillary zone electrophoresis, nuclear magnetic resonance spectroscopy, fluorescence spectroscopy, and combinations thereof.

One skilled in the art may use one of several methods to screen chemical entities or fragments for their ability to associate with an *S. aureus* MurB or *S. aureus* MurB-like binding pocket. This process may begin by visual inspection of, for example, an *S. aureus* MurB or *S. aureus* MurB-like binding pocket on the computer screen based on the *S. aureus* MurB structure coordinates listed in Figure 4 or other coordinates which define a similar shape generated from the machine-readable storage medium. Selected fragments or chemical entities may then be positioned in a variety of orientations, or docked, within the binding pocket. Docking may be accomplished using software such as QUANTA and SYBYL, followed by energy minimization and molecular dynamics with standard molecular mechanics forcefields, such as CHARMM and AMBER.

Specialized computer programs may also assist in the process of selecting fragments or chemical entities. Examples include GRID (P.J. Goodford, J. Med. Chem. 28:849-857 (1985); available from Oxford University, Oxford, UK); MCSS (A. Miranker et al., Proteins: Struct. Funct. Gen.,11:29-34 (1991); available from Molecular Simulations, San Diego, CA); AUTODOCK (D.S. Goodsell et al., Proteins: Struct. Funct. Genet. 8:195-202 (1990); available from Scripps Research Institute, La Jolla, CA); and DOCK (I.D. Kuntz et al., J. Mol. Biol. 161:269-288 (1982); available from University of California, San Francisco, CA).

Once suitable chemical entities or fragments have been selected, they can be assembled into a single compound or complex. Assembly may be preceded by visual inspection of the relationship of the fragments to each other on the three-dimensional image displayed on a computer screen in relation to the structure coordinates of *S. aureus* MurB. This would be followed by manual model building using software such as QUANTA or SYBYL (Tripos Associates, St. Louis, MO).

Useful programs to aid one of skill in the art in connecting the individual chemical entities or fragments include, without limitation, CAVEAT (P.A. Bartlett et al., in Molecular Recognition in Chemical and Biological Problems," Special Publ., Royal Chem. Soc., 78:182-196 (1989); G. Lauri et al., J. Comput. Aided Mol. Des. 8:51-66 (1994); available from the University of California, Berkeley, CA); 3D database systems such as ISIS (available from MDL Information Systems, San Leandro, CA; reviewed in Y.C. Martin, J. Med. Chem. 35:2145-2154 (1992)); and HOOK (M.B. Eisen et al., Proteins: Struc., Funct., Genet. 19:199-221 (1994); available from Molecular Simulations, San Diego, CA).

S. aureus MurB binding compounds may be designed "*de novo*" using either an empty binding site or optionally including some portion(s) of a known inhibitor(s). There are many *de novo* ligand design methods including, without limitation, LUDI (H.-J. Bohm, J. Comp. Aid. Molec. Design. 6:61-78 (1992); available from Molecular Simulations Inc., San Diego, CA); LEGEND (Y. Nishibata et al., Tetrahedron, 47:8985 (1991); available from Molecular Simulations Inc., San Diego, CA); LeapFrog (available from Tripos Associates, St. Louis, MO); and SPROUT (V. Gillet et al., J. Comput. Aided Mol. Design 7:127-153 (1993); available from the University of Leeds, UK).

Once a compound has been designed or selected by the above methods, the efficiency with which that entity may bind to or interfere with an *S. aureus* MurB or *S. aureus* MurB-like binding pocket may be tested and optimized by computational evaluation. For example, an effective *S. aureus* MurB or *S. aureus* MurB-like binding pocket inhibitor must preferably demonstrate a relatively small difference in energy between its bound and free states (i.e., a small deformation energy of binding). Thus, the most efficient *S.*

aureus MurB or *S. aureus* MurB-like binding pocket inhibitors should preferably be designed with a deformation energy of binding of not greater than about 10 kcal/mole; more preferably, not greater than 7 kcal/mole. *S. aureus* MurB or *S. aureus* MurB-like binding pocket inhibitors may interact with the binding pocket in more than one conformation that is similar in overall binding energy. In those cases, the deformation energy of binding is taken to be the difference between the energy of the free entity and the average energy of the conformations observed when the inhibitor binds to the protein.

An entity designed or selected as binding to or interfering with an *S. aureus* MurB or *S. aureus* MurB-like binding pocket may be further computationally optimized so that in its bound state it would preferably lack repulsive electrostatic interaction with the target enzyme and with the surrounding water molecules. Such non-complementary electrostatic interactions include repulsive charge-charge, dipole-dipole, and charge-dipole interactions.

Specific computer software is available in the art to evaluate compound deformation energy and electrostatic interactions. Examples of programs designed for such uses include: Gaussian 94, revision C (M.J. Frisch, Gaussian, Inc., Pittsburgh, PA 81995); AMBER, version 4.1 (P.A. Kollman, University of California at San Francisco, 81995); QUANTA/CHARMM (Molecular Simulations, Inc., San Diego, CA 81995); Insight II/Discover (Molecular Simulations, Inc., San Diego, CA 81995); DelPhi (Molecular Simulations, Inc., San Diego, CA 81995); and AMSOL (Quantum Chemistry Program Exchange, Indiana University). These programs may be implemented, for instance, using a Silicon Graphics workstation such as an Indigo^2 with "IMPACT" graphics. Other hardware systems and software packages will be known to those skilled in the art.

Another approach encompassed by this invention is the computational screening of small molecule databases for chemical entities or compounds that can bind in whole, or in part, to a *S. aureus* MurB or *S. aureus* MurB-like binding pocket. In this screening, the quality of fit of such entities to the binding site may be judged either by shape complementarity or by estimated interaction energy (E.C. Meng et al., J. Comp. Chem., 13, pp. 505-524 (1992)).

This invention also enables the development of chemical entities that can isomerize to short-lived reaction intermediates in the chemical reaction of a substrate or other compound that interferes with or with *S. aureus* MurB. Time-dependent analysis of structural changes in *S. aureus* MurB during its interaction with other molecules is carried out. The reaction intermediates of *S. aureus* MurB can also be deduced from the reaction product in co-complex with *S. aureus* MurB. Such information is useful to design improved analogs of known *S. aureus* MurB inhibitors or to design novel classes of inhibitors based on the reaction intermediates of the *S. aureus* MurB and inhibitor co-complex. This provides a novel route for designing *S. aureus* MurB inhibitors with both high specificity and stability.

Yet another approach to rational drug design involves probing the *S. aureus* MurB crystal of the invention with molecules comprising a variety of different functional groups to determine optimal sites for interaction between candidate *S. aureus* MurB inhibitors and the protein. For example, high resolution x-ray diffraction data collected from crystals soaked in or co-crystallized with other molecules allows the determination of where each type of solvent molecule sticks. Molecules that bind tightly to those sites can then be further modified and synthesized and tested for their MurB inhibitor activity (J. Travis, Science, 262:1374 (1993)).

In a related approach, iterative drug design is used to identify inhibitors of *S. aureus* MurB. Iterative drug design is a method for optimizing associations between a protein and a compound by determining and evaluating the three-dimensional structures of successive sets of protein/compound complexes. In iterative drug design, crystals of a series of protein/compound complexes are obtained and then the three-dimensional structures of each complex is solved. Such an approach provides insight into the association between the proteins and compounds of each complex. This is accomplished by selecting compounds with inhibitory activity, obtaining crystals of this new protein/compound complex, solving the three dimensional structure of the complex, and comparing the associations between the new protein/compound complex and previously solved protein/compound complexes. By observing how

changes in the compound affected the protein/compound associations, these associations may be optimized.

A compound that is identified or designed as a result of any of these methods can be obtained (or synthesized) and tested for its biological activity, e.g., inhibition of MurB activity.

Pharmaceutical Compositions

Pharmaceutical compositions of this invention comprise an inhibitor of *S. aureus* MurB activity identified according to the invention, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier, adjuvant, or vehicle. The term "pharmaceutically acceptable carrier" refers to a carrier(s) that is "acceptable" in the sense of being compatible with the other ingredients of a composition and not deleterious to the recipient thereof. Optionally, the pH of the formulation is adjusted with pharmaceutically acceptable acids, bases, or buffers to enhance the stability of the formulated compound or its delivery form.

Methods of making and using such pharmaceutical compositions are also included in the invention. The pharmaceutical compositions of the invention can be administered orally, parenterally, by inhalation spray, topically, rectally, nasally, buccally, vaginally, or via an implanted reservoir. Oral administration or administration by injection is preferred. The term parenteral as used herein includes subcutaneous, intracutaneous, intravenous, intramuscular, intra-articular, intrasynovial, intrasternal, intrathecal, intralesional, and intracranial injection or infusion techniques.

Dosage levels of between about 0.01 and about 100 mg/kg body weight per day, preferably between about 0.5 and about 75 mg/kg body weight per day of the *S. aureus* MurB inhibitory compounds described herein are useful for the prevention and treatment of *S. aureus* MurB mediated disease. Typically, the pharmaceutical compositions of this invention will be administered from about 1 to about 5 times per day or alternatively, as a continuous infusion. Such administration can be used as a chronic or acute therapy. The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode

of administration. A typical preparation will contain from about 5% to about 95% active compound (w/w). Preferably, such preparations contain from about 20% to about 80% active compound.

In order that this invention be more fully understood, the following examples are set forth. These examples are for the purpose of illustration only and are not to be construed as limiting the scope of the invention in any way.

EXAMPLES

Example 1: Analysis of the Structure of *S. aureus* MurB

Expression of MurB and Incorporation of Selenomethionine

S. aureus MurB was expressed using UC 15169, *E. coli* construct K12S (F' *lacI*^{q}) (pQE-10 murB). Genes and polypeptides derived from *S. aureus*, including *S. aureus* and MurB, are published in EP 786519 A2 and WO 0012678, both assigned to Human Genome Sciences. MurB cloned into pQE-10 (Qiagen) was obtained from Human Genome Sciences. For expression, the plasmid was transformed into the *E. coli* K12S F' cell line which has an ampicillin resistance marker. Stock supplies of the culture were maintained at -80°C in Luria Broth containing ampicillin at 100 μg/mL with 10% glycerol added as a cryopreservative agent.

Seed fermentations were prepared in 100 mL volumes of M9 medium contained in 500 mL wide mouth fermentation flasks. The formulation of basal M9 utilized for these studies was Na_2HPO_4, 6 g; KH_2PO_4, 3 g; NH_4Cl, 1.0 g; and NaCl, 0.5 g per liter of deionized water. The pH was adjusted to 7.4 with concentrated KOH. The medium was sterilized by autoclaving for 30 minutes. Prior to inoculation, the following filter sterilized solutions were added per liter of basal medium: 1M $MgSO_4$, 1.0 mL; 1M $CaCl_2$, 0.3 mL; trace metal salts solution, 0.3 mL and 20% glucose, 20 mL. The trace metal salts solution contained per liter of deionized water: $MgCl_2 \cdot 6H_2O$, 39.44 g; $MnSO_4 \cdot H_2O$, 5.58 g; $FeSO_4 \cdot 7H_2O$, 1.11 g; $Na_2MoO_4 \cdot 2H_2O$, 0.48 g; $CaCl_2$, 0.33 g; NaCl, 0.12 g; and ascorbic acid, 1.0 g. Filter sterilized ampicillin was

added to the medium at a final concentration of 100 μg/mL. A 0.1 mL aliquot of the stock culture was inoculated into the medium and allowed to grow at 37 °C for 18 -20 hours with a shaking rate of 200 rpm. The mature seed culture was harvested by centrifugation and then resuspended in an equal volume of M9 medium. The resuspended seed was used to inoculate expression fermentations at a rate of 3%.

For expression of selenomethionine MurB, M9 media was again utilized in 100 mL volumes containing 100 μg/mL of ampicillin. Multiple flasks were employed to achieve the desired production volume. Since UC 15169 is not a methionine auxotroph, incorporation of selenomethionine was accomplished through down-regulation of methionine biosynthesis just prior to induction of MurB expression with IPTG, isopropyl thio-β-D-galactosidase (G.D. Van Duyne et al., J. Mol. Biol. 229, 105-24 (1993); T.E. Benson et al., Nat. Struct. Biol. 2, 644-53 (1995)). The culture was grown at 37°C with a shaking rate of 200 rpm until an A_{600} of ~0.6. At this point, the following filter sterilized amino acids were added. L-lysine, L-threonine, and L-phenylalanine were added to final concentrations of 100 μg/mL. L-leucine, L-isoleucine, and L-valine were added to final concentrations of 50 μg/mL. Filter sterilized L-selenomethionine was added simultaneously to a final concentration of 50 μg/mL. After 15 - 20 minutes, protein expression was induced by the addition of filter sterilized IPTG to a final concentration of 1 mM. Growth of the culture was continued at 200 rpm for an additional 4 hours until an A_{600} of ~2.0. This coincided with maximum growth and maximum expression of MurB. Cells were then harvested by centrifugation and frozen at -80°C. Under these conditions, the average yield of cell paste was 4 - 4.5 g/L. Selenomethionyl MurB comprised roughly 2 - 5% of the total cell protein with > 75% expressed in the soluble form.

Purification of Selenomethionine MurB

All operations were performed at 4°C and 2-mercaptoethanol and DTT were added to buffers immediately before use. Three hundred milliliters of equilibration buffer (50 mM Tris, pH 7.8, 500 mM NaCl, 10% glycerol, 25 mM

imidazole, 5 mM 2-mercaptoethanol) containing 0.2 mg/mL DNAse I (Boehringer Mannheim #104159) was added to 26 g of cell paste obtained from 6 L of fermentation broth and was resuspended by using a Tekmar Tissumizer set on a power setting of 60. The suspension was homogenized by passing it twice through a Rannie homogenizer at 10,000 PSI. The homogenate was centrifuged at 39,200 x g for 60 minutes in a JA20 rotor in a Beckman J2-21 centrifuge. The supernatant was filtered by using a Nalgene 0.2 μm CN filter unit and applied to a Qiagen NTA Superflow column charged with nickel (column volume of 7.9 mL). The column was then washed with 4 column volumes of equilibration buffer and 22 column volumes of wash buffer (50 mM Tris, pH 7.8, 500 mM NaCl, 10% glycerol, 50 mM imidazole, 5 mM 2-mercaptoethanol) at a flow rate of 108 mL/hr and eluted with 2.5 column volumes of elution buffer (50 mM Tris, pH 7.8, 500 mM NaCl, 10% glycerol, 300 mM imidazole, 5 mM 2-mercaptoethanol) at a flow rate of 60 mL/hr. DTT was added to the eluted material to a final concentration of 10 mM and the treated material was dialyzed for 22 hours against two changes of nitrogen sparged dialysis buffer (50 mM Tris pH 7.8, 500 mM NaCl, 10% glycerol, and 10 mM DTT). After dialysis the sample was sterile filtered, fractionated, and stored at -80°C.

The protein concentration was 2.42 mg/mL as determined by amino acid analysis. The prepared MurB protein had the correct N-terminal sequence for the first 20 residues. The mass, as measured by electrospray mass spectrometry, was 36,220 Da, in excellent agreement with the theoretical mass of 36,207 Da, indicating full incorporation of the five selenomethionines into the protein. Amino acid analysis gave a correlation coefficient of 0.99 between the recovered and theoretical amino acid composition, indicative not only of high purity but also of the correct amino acid composition in the protein.

Protein Crystallization

Protein samples were buffer exchanged into 20mM HEPES pH 7.5, 5mM 2-mercaptoethanol and concentrated to 20 mg/mL using an Ultrafree 0.5 centrifugal filters with a Biomax 10K membrane (Millipore, Bedford, MA). Selenomethionine MurB crystals were grown in 3μL + 3μL sitting drops in

9.75% PEG 8000, 0.1 M cacodylic acid pH 6.5, 0.55 M ammonium sulfate, 20% DMSO, 5mM 2-mercaptoethanol with 1mM EP-UDPGlcNAc substrate. These conditions were originally identified by screening for crystallization conditions with the methionine incorporated *S. aureus* MurB. The hexagonal shaped crystals grew over a period of two to three weeks. The mother liquor served as the cryoprotectant for freezing during data collection at 100 K in liquid nitrogen.

Data Collection and Structure Determination.

Access to synchrotron radiation at the Advance Photon Source at Argonne National Labs (IMCA-CAT, Beamline 17-ID) afforded the opportunity to solve the *S. aureus* MurB structure by multiple anomalous dispersion (MAD) phasing. EXAFS analysis revealed a sharp selenium K edge for the selenomethionine MurB (data not shown). A three wavelength experiment was carried out with a low energy wavelength (12,000 eV, 1.0332 Å), a wavelength corresponding to the inflection point of the absorption edge (12,659.4 eV, 0.97939 Å), and a wavelength collected at the peak of the absorption edge (12,660.8 eV, 0.97928 Å). All diffraction data were collected on a 2k by 2k Brüker CCD detector.

Data sets at each wavelength were processed separately with the program SAINT (Siemens Analytical X-ray Systems, Madison, WI) while keeping the anomalous pairs separate (Table 7). The inflection point and peak data sets were scaled to the remote energy data set using SCALEIT in CCP4 (Collaborative Computational Project, N.4 Acta Cryst. D50, 760-63 (1994)) by treating the remote wavelength as native. Anomalous and dispersive difference Patterson maps showed strong signals for 4 of the 5 selenium atoms suggesting the N-terminal methionine was disordered. Locations of the selenium sites were determined using the automated Patterson solution routine in SHELX (G.M. Sheldrick & R.O. Gould, Acta Cryst. B51, 423-31 (1995)). The location of each selenium site was confirmed by the ability of individual sites to generate phases which could identify the other sites in cross difference Fourier calculations. All heavy atom parameter refinement and phasing calculations were carried out with MLPHARE (Z. Otwinowski, *in* Isomorphous Replacement and Anomalous

Scattering 80-86 (W. Wolf et al., eds., SERC Daresbury Laboratory, Warrington) (1991); Collaborative Computational Project, N.4 Acta Cryst. D50, 760-63 (1994)) by treating the remote wavelength as native and the edge and peak wavelengths as derivatives (V. Ramakrishnan et al., Nature 362, 219-23 (1993)). The phases were subsequently subjected to solvent flattening using the program DM (K.D. Cowtan & P. Main, Acta Cryst. D49, 148-57 (1993); K.D. Cowtan & P. Main, Acta Cryst. D54, 487-93 (1998); Collaborative Computational Project, N.4 Acta Cryst. D50, 760-63 (1994)).

Model building was performed using the program CHAIN (J.S. Sack, J. Mol. Graph. 6, 224-25 (1988)). The *E. coli* MurB model was used as a template for model building in order to speed the placement of the main chain atoms. All refinement steps were carried out using XPLOR 3.8.5.1 and XPLOR 98.0 (A.T. Brunger, Methods. Mol. Biol. 56, 245-66 (1996)) against the 1.0332Å (low energy) data set. Several rounds of torsional dynamics (L.M. Rice & A.T. Brunger, Proteins 19, 277-90 (1994)) and simulated annealing (A.T. Brunger, J. Mol. Biol. 203, 803-16 (1988)) with rebuilding after each round of refinement were carried out. In the later stages of refinement, a bulk solvent model was included in order to properly account for inclusion of lower resolution data (J.S. Jiang & A.T. Brunger, J. Mol. Biol. 243, 100-15 (1994)). Progress of the refinement was monitored by the Free R factor which was calculated for 10% of the reflections that were not included in refinement (A.T. Brunger, Nature 355, 472-75 (1992)). Analysis by PROCHECK showed good main chain geometry and side chain torsion angles (R.A. Laskowski et al., J. Appl. Cryst. 26, 283-91 (1993)). Figure 3 was made with Setor (S.V. Evans, J. Mol. Graph. 11, 134-38 (1993)). Figure 5 was made with Molscript 2.1 (P. Kraulis, J. Appl. Cryst. 24, 946-50 (1991)) and Raster3D (E.A. Merritt & D.J. Bacon, Meth. Enzymol. 277, 505-24 (1997)), and Figures 7 and 11 were made with Molscript 2.1 only.

Sequence Listing Free Text

SEQ ID NO:1	Recombinant *S. aureus* MurB protein including polyhistidine (His_6) region.
SEQ ID NO:2	*E. Coli* MurB protein.
SEQ ID NO:3	*Helicobacter pylori* MurB protein.
SEQ ID NO:4	*Aquifex aeolicus* MurB protein.
SEQ ID NO:5	*Bacillus subtilis* MurB protein.
SEQ ID NO:6	*Borrelia burgdorferi* MurB protein.
SEQ ID NO:7	*Chlamydia pneumoniae* MurB protein.
SEQ ID NO:8	*Rickettsia prowazekii* MurB protein.
SEQ ID NO:9	*Haemophilus influenzae* MurB protein.
SEQ ID NO:10	*Salmonella typhimurium* MurB protein.
SEQ ID NO:11	*Bordetella pertussis* MurB protein.